CLAIMS

1. A decoder for an automatic speech recognition system for determining one or more candidate text unit concatenations according to a predetermined criterion and which correspond to a speech segment, the decoder comprising:

a processor arranged to receive a sequence of feature vectors corresponding to the speech segment;

the processor arranged to map with different likelihood values the feature vectors to sequences of nodes in a decoding network, every sequence representing a concatenation of text units;

the processor arranged to determine one or more candidate node sequences in the decoding network corresponding to the candidate text unit concatenations by implementing a dynamic programming token passing algorithm in which each token corresponds to a node and is associated with a number of text unit concatenations and likelihood values for these concatenations, and wherein a token associated with a node in the decoding network is derived from the tokens associated with the previous nodes in the network;

wherein tokens from different nodes that are to be passed to a common node are combined to generate a new token corresponding to the common node and associated with an identifier for text unit concatenations and likelihood values associated with the previous tokens of said different nodes.

2. A decoder according to claim 1 wherein the processor is further arranged to merge a said token having a said identifier, the text unit concatenations of the said previous tokens being associated with said merged token dependent on their corresponding likelihood values.

3. A decoder according to claim 2 wherein said merging is only delayed if the token has a likelihood value below a delay-merge threshold.

4. A decoder according to claim 1 wherein the processor is further arranged to prune tokens having likelihood values below a prune threshold.

5. A decoder according to claim 1 wherein the plurality of candidate text unit concatenations in a token are the text unit concatenations with the N-best likelihood values.

6. A decoder according to claim 1 wherein the tokens are additionally associated with a number of category markers each corresponding to a said text unit concatenation, each category marker being associated with one of a plurality of categories, such that the plurality of candidate text unit concatenations in a token are the text unit concatenations with the best likelihood values in said plurality of categories.

7. A decoder according to claim 6 wherein the plurality of candidate text unit concatenations are the text unit concatenations with the N-best likelihood values in each said category.

8. A decoder according to claim 1 wherein the multiple text unit concatenations associated with each token are used to allow a statistical language model score to be added to the likelihood values associated with said text unit concatenations.

9. A decoder according to claim 1 wherein the dynamic programming token passing algorithm is a Viterbi algorithm.

10. A decoder according claim 1 wherein the dynamic programming token passing algorithm is a Baum-Welch algorithm.

11. A decoder according to claim 1 wherein the tokens and the text unit concatenations are stored in logically separate memories, and wherein a logically separate list data-structure is used to associate tokens with their text unit concatenations or identifiers and corresponding likelihood values.

12. A decoder according to claim 1 wherein the mapping means comprises an acoustic model based on Hidden Markov Models.

13. An automatic speech recognition system comprising a decoder according to claim 1.

14. A voice activated control or navigation system for in car use, the system comprising an automatic speech recognition system according to claim 13.

15. A method of decoding for determining a plurality of candidate text unit concatenations according to a predetermined criterion and corresponding to a speech segment in an automatic speech recognition system, the method comprising:

receiving a sequence of feature vectors corresponding to the speech segment;

mapping with different likelihood values the feature vectors to sequences of nodes in a decoding network, every sequence representing a concatenation of text units;

determining one or more candidate node sequences in the decoding network corresponding to the candidate text unit concatenations by implementing a dynamic programming token passing algorithm in which each token corresponds to a node and is associated with a number of text unit concatenations and likelihood values for these concatenations, and wherein a token associated with a node in the decoding network is derived from the tokens associated with the previous nodes in the network;

wherein tokens from different nodes that are to be passed to a common node are combined to generate a new token corresponding to the common node and associated with an identifier for text unit concatenations and likelihood values associated with the previous tokens of said different nodes.

16. A method according to claim 15 further comprising merging a said token having a said identifier, the text unit concatenations of the said previous tokens being associated with said merged token dependent on their corresponding likelihood values.

17. A method according to claim 16 wherein said merging is only delayed if the token has a likelihood value below a delay-merge threshold.

18. A method according to claim 15 further comprising pruning tokens having likelihood values below a prune threshold.

19. A method according to claim 15 wherein the plurality of candidate text unit concatenations in a token are the text unit concatenations with the N-best likelihood values.

20. A method according to claim 15 wherein the tokens are additionally associated with a number of category markers each corresponding to a said text unit concatenation, each category marker being associated with one of a plurality of categories, such that the plurality of candidate text unit concatenations in a token are the text unit concatenations with the best likelihood values in said plurality of categories.

21. A method according to claim 20 wherein the plurality of candidate text unit concatenations are the text unit concatenations with the N-best likelihood values in each said category.

22. A method according to claim 15 wherein the multiple text unit concatenations associated with each token are used to allow a statistical language model score to be added to the likelihood values associated with said text unit concatenations.

23. A method according to claim 15 wherein the dynamic programming token passing algorithm is a Viterbi algorithm.

24. A method according to claim 15 wherein the dynamic programming token passing algorithm is a Baum-Welch algorithm.

25. A method according to claim 15 wherein the tokens and the text unit concatenations are stored in logically separate memories, and wherein a logically separate list data-structure is used to associate tokens with their text unit concatenations or identifiers and corresponding likelihood values.

26. A method of automatically recognising speech comprising a method of decoding according to claim 15.

27. A carrier medium carrying processor control code for implementing a method according to claim 15.